

1600

CRF Errors Corrected by the STIC Systems Branch.

CRF Processing Date: 8/5/2003

Edited by: M

Verified by: M

(STIC staff)

Serial Number: 09/689,343E

ENTERED

RECEIVED #22

AUG 11 2003

TECH CENTER 1600/2900

- ☐ Changed a file from non-ASCII to ASCII
- ☐ Changed the margins in cases where the sequence text was "wrapped" down to the next line.
- ☐ Edited a format error in the Current Application Data section, specifically: \_\_\_\_\_
- ☐ Edited the Current Application Data section with the actual current number. The number input by the applicant was ☐ the prior application data; or ☐ other \_\_\_\_\_
- ☐ Added the mandatory heading and subheadings for "Current Application Data".
- ☐ Edited the "Number of Sequences" field. The applicant spelled out a number instead of using an integer.
- ☐ Changed the spelling of a mandatory field (the headings or subheadings), specifically: \_\_\_\_\_
- ☐ Corrected the SEQ ID NO when obviously incorrect. The sequence numbers that were edited were: \_\_\_\_\_
- ☐ Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited: \_\_\_\_\_
- ☐ Corrected subheading placement. All responses must be on the same line as each subheading. If the applicant placed a response below the subheading, this was moved to its appropriate place.
- ☐ Inserted colons after headings/subheadings. Headings edited included: \_\_\_\_\_
- ☐ Deleted extra, invalid, headings used by an applicant, specifically: \_\_\_\_\_
- ☐ Deleted: ☐ non-ASCII "garbage" at the beginning/end of files; ☐ secretary initials/filename at end of file; ☐ page numbers throughout text; ☐ other invalid text, such as \_\_\_\_\_
- ☐ Inserted mandatory headings, specifically: \_\_\_\_\_
- ☐ Corrected an obvious error in the response, specifically: \_\_\_\_\_
- ☐ Edited identifiers where upper case is used but lower case is required, or vice versa.
- ☐ Corrected an error in the Number of Sequences field, specifically: \_\_\_\_\_
- ☐ A "Hard Page Break" code was inserted by the applicant. All occurrences had to be deleted.
- ☐ Deleted ending stop codon in amino acid sequences and adjusted the "(A)Length:" field accordingly (error due to a PatentIn bug). Sequences corrected: \_\_\_\_\_
- ☒ Other: inserted hard returns in seqs 3 and 11

\*Examiner: The above corrections must be communicated to the applicant in the first Office Action. DO NOT send a copy of this form.

3/1/95



1600

## RAW SEQUENCE LISTING

DATE: 08/05/2003

PATENT APPLICATION: US/09/689,343E

TIME: 18:28:00

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\08052003\I689343E.raw

3 <110> APPLICANT: Vaisvila, Romualdus  
 4 Morgan, Richard D.  
 5 Kucera, Rebecca B.  
 6 Claus, Toby B.  
 7 Raleigh, Elisabeth A.  
 9 <120> TITLE OF INVENTION: Method For Cloning And Producing The MseI Restriction  
 10 Endonuclease  
 12 <130> FILE REFERENCE: NEB-181  
 14 <140> CURRENT APPLICATION NUMBER: US 09/689,343E  
 15 <141> CURRENT FILING DATE: 2000-10-12  
 17 <160> NUMBER OF SEQ ID NOS: 21  
 19 <170> SOFTWARE: PatentIn version 3.1  
 21 <210> SEQ ID NO: 1  
 22 <211> LENGTH: 903  
 23 <212> TYPE: DNA  
 24 <213> ORGANISM: Micrococcus sp.  
 26 <220> FEATURE:  
 27 <221> NAME/KEY: CDS  
 28 <222> LOCATION: (1)..(900)  
 29 <223> OTHER INFORMATION:

W--&gt; 31 &lt;400&gt; 1

33 atg cct atc tcg acc gtc tgg acg ccg gac gga gac gac ctc atc gtg 48  
 34 Met Pro Ile Ser Thr Val Trp Thr Pro Asp Gly Asp Asp Leu Ile Val  
 35 1 5 10 15  
 37 gag gcg gac aac ctc gat ttc att caa acg ctc ccc gac gcg agc ttc 96  
 38 Glu Ala Asp Asn Leu Asp Phe Ile Gln Thr Leu Pro Asp Ala Ser Phe  
 39 20 25 30  
 41 cga atg atc tac atc gat ccg ccg ttc aac aca ggg cga acg cag cgg 144  
 42 Arg Met Ile Tyr Ile Asp Pro Pro Phe Asn Thr Gly Arg Thr Gln Arg  
 43 35 40 45  
 45 ctt cag tcg ctc aag acg acc cgc tcg gtc aca ggg tcg cga gtc ggc 192  
 46 Leu Gln Ser Leu Lys Thr Thr Arg Ser Val Thr Gly Ser Arg Val Gly  
 47 50 55 60  
 49 ttc aaa ggc cag acg tac gac acg gtc aag agc act ctg cac tcg tat 240  
 50 Phe Lys Gly Gln Thr Tyr Asp Thr Val Lys Ser Thr Leu His Ser Tyr  
 51 65 70 75 80  
 53 gac gac gct ttc acc gac tat tgg tcg ttc ctc gaa ccg cgt ctc ctg 288  
 54 Asp Asp Ala Phe Thr Asp Tyr Trp Ser Phe Leu Glu Pro Arg Leu Leu  
 55 85 90 95  
 57 gag gct tgg cgg ttg ctc acc cct gac ggc gcg ctc tat ctt cat ctg 336  
 58 Glu Ala Trp Arg Leu Leu Thr Pro Asp Gly Ala Leu Tyr Leu His Leu  
 59 100 105 110  
 60 gat tac cgc gag gtt cac tac gcc aag gtc gtc ctc gac gcg atg ttc 384

## RAW SEQUENCE LISTING

DATE: 08/05/2003

PATENT APPLICATION: US/09/689,343E

TIME: 18:28:00

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\08052003\I689343E.raw

```

61 Asp Tyr Arg Glu Val His Tyr Ala Lys Val Val Leu Asp Ala Met Phe
62      115      120      125
64 gga cgc gaa agc ttc ctg aac gag ctg atc tgg gcg tac gac tac ggc      432
65 Gly Arg Glu Ser Phe Leu Asn Glu Leu Ile Trp Ala Tyr Asp Tyr Gly
66      130      135      140
68 gcg cgc tcg aag agc aag tgg ccc acc aag cac gac aac atc ctc gtg      480
69 Ala Arg Ser Lys Ser Lys Trp Pro Thr Lys His Asp Asn Ile Leu Val
70 145      150      155      160
72 tat gtg aag gac ccg aac aac tac gtc tgg aac ggt cag gat gta gat      528
73 Tyr Val Lys Asp Pro Asn Asn Tyr Val Trp Asn Gly Gln Asp Val Asp
74      165      170      175
76 cgc gag ccc tac atg gcg ccc ggg ctc gtt aca ccc gag aag gta gcg      576
77 Arg Glu Pro Tyr Met Ala Pro Gly Leu Val Thr Pro Glu Lys Val Ala
78      180      185      190
80 ctt ggc aag ctg ccc acc gac gtc tgg tgg cac aca atc gtt ccg cct      624
81 Leu Gly Lys Leu Pro Thr Asp Val Trp Trp His Thr Ile Val Pro Pro
82      195      200      205
84 gcg agc aaa gag cgc acc ggg tac gcg aca cag aag ccg gtc ggc atc      672
85 Ala Ser Lys Glu Arg Thr Gly Tyr Ala Thr Gln Lys Pro Val Gly Ile
86      210      215      220
88 atc cgt cgc atg att cag gcg agc agc aat gaa ggc gac tgg gtt ctg      720
89 Ile Arg Arg Met Ile Gln Ala Ser Ser Asn Glu Gly Asp Trp Val Leu
90 225      230      235      240
92 gat ttc ttc gct ggt agt ggg acg acc ggc gcc gcg gcc cgc cag ctc      768
93 Asp Phe Phe Ala Gly Ser Gly Thr Thr Gly Ala Ala Ala Arg Gln Leu
94      245      250      255
96 gga cgc cgt ttt gtg ctc gta gac gtc aac cca gaa gca atc gcg gta      816
97 Gly Arg Arg Phe Val Leu Val Asp Val Asn Pro Glu Ala Ile Ala Val
98      260      265      270
100 atg gca aaa cgg ttg gat gac ggg gca ttg gac acc agc gtg acg atc      864
101 Met Ala Lys Arg Leu Asp Asp Gly Ala Leu Asp Thr Ser Val Thr Ile
102      275      280      285
104 gtg cag act ccc cag agt gac cca cga acc gac gga tga      903
105 Val Gln Thr Pro Gln Ser Asp Pro Arg Thr Asp Gly
106      290      295      300
108 <210> SEQ ID NO: 2
109 <211> LENGTH: 300
110 <212> TYPE: PRT
111 <213> ORGANISM: Micrococcus sp.
113 <400> SEQUENCE: 2
115 Met Pro Ile Ser Thr Val Trp Thr Pro Asp Gly Asp Asp Leu Ile Val
116 1      5      10      15
118 Glu Ala Asp Asn Leu Asp Phe Ile Gln Thr Leu Pro Asp Ala Ser Phe
119      20      25      30
121 Arg Met Ile Tyr Ile Asp Pro Pro Phe Asn Thr Gly Arg Thr Gln Arg
122      35      40      45
124 Leu Gln Ser Leu Lys Thr Thr Arg Ser Val Thr Gly Ser Arg Val Gly
125      50      55      60
127 Phe Lys Gly Gln Thr Tyr Asp Thr Val Lys Ser Thr Leu His Ser Tyr

```

## RAW SEQUENCE LISTING

DATE: 08/05/2003

PATENT APPLICATION: US/09/689,343E

TIME: 18:28:00

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\08052003\I689343E.raw

```

128 65          70          75          80
130 Asp Asp Ala Phe Thr Asp Tyr Trp Ser Phe Leu Glu Pro Arg Leu Leu
131          85          90          95
133 Glu Ala Trp Arg Leu Leu Thr Pro Asp Gly Ala Leu Tyr Leu His Leu
134          100          105          110
136 Asp Tyr Arg Glu Val His Tyr Ala Lys Val Val Leu Asp Ala Met Phe
137          115          120          125
139 Gly Arg Glu Ser Phe Leu Asn Glu Leu Ile Trp Ala Tyr Asp Tyr Gly
140          130          135          140
142 Ala Arg Ser Lys Ser Lys Trp Pro Thr Lys His Asp Asn Ile Leu Val
143 145          150          155          160
145 Tyr Val Lys Asp Pro Asn Asn Tyr Val Trp Asn Gly Gln Asp Val Asp
146          165          170          175
148 Arg Glu Pro Tyr Met Ala Pro Gly Leu Val Thr Pro Glu Lys Val Ala
149          180          185          190
151 Leu Gly Lys Leu Pro Thr Asp Val Trp Trp His Thr Ile Val Pro Pro
152          195          200          205
154 Ala Ser Lys Glu Arg Thr Gly Tyr Ala Thr Gln Lys Pro Val Gly Ile
155          210          215          220
157 Ile Arg Arg Met Ile Gln Ala Ser Ser Asn Glu Gly Asp Trp Val Leu
158 225          230          235          240
160 Asp Phe Phe Ala Gly Ser Gly Thr Thr Gly Ala Ala Ala Arg Gln Leu
161          245          250          255
163 Gly Arg Arg Phe Val Leu Val Asp Val Asn Pro Glu Ala Ile Ala Val
164          260          265          270
166 Met Ala Lys Arg Leu Asp Asp Gly Ala Leu Asp Thr Ser Val Thr Ile
167          275          280          285
169 Val Gln Thr Pro Gln Ser Asp Pro Arg Thr Asp Gly
170          290          295          300

```

172 &lt;210&gt; SEQ ID NO: 3

173 &lt;211&gt; LENGTH: 1236

174 &lt;212&gt; TYPE: DNA

175 &lt;213&gt; ORGANISM: Unknown

177 &lt;220&gt; FEATURE:

178 &lt;223&gt; OTHER INFORMATION: Environmental DNA

180 &lt;220&gt; FEATURE:

181 &lt;221&gt; NAME/KEY: CDS

182 &lt;222&gt; LOCATION: (1)..(1233)

183 &lt;223&gt; OTHER INFORMATION:

185 &lt;220&gt; FEATURE:

186 &lt;221&gt; NAME/KEY: misc\_feature

187 &lt;222&gt; LOCATION: (198)..(198)

188 &lt;223&gt; OTHER INFORMATION: Xaa = any amino acid

W--&gt; 190 &lt;220&gt;

191 &lt;221&gt; NAME/KEY: misc\_feature

192 &lt;222&gt; LOCATION: (594)..(594)

193 &lt;223&gt; OTHER INFORMATION: N= G, A, C or T

195 &lt;400&gt; SEQUENCE: 3

197 atg cct aca ctg gat tgg ccc ggt aaa cag tta agc ttc cca cca gct

48

## RAW SEQUENCE LISTING

DATE: 08/05/2003

PATENT APPLICATION: US/09/689,343E

TIME: 18:28:00

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\08052003\I689343E.raw

198	Met	Pro	Thr	Leu	Asp	Trp	Pro	Gly	Lys	Gln	Leu	Ser	Phe	Pro	Pro	Ala	
199	1				5					10					15		
201	acc	tcc	ttg	cat	ctg	gag	agt	gtg	gtc	act	gag	gga	gcg	gag	tca	ccg	96
202	Thr	Ser	Leu	His	Leu	Glu	Ser	Val	Val	Thr	Glu	Gly	Ala	Glu	Ser	Pro	
203				20					25					30			
205	cct	aat	cgt	ctg	att	tgg	gcg	gac	aac	ctg	ccg	cta	atg	gta	gat	ttg	144
206	Pro	Asn	Arg	Leu	Ile	Trp	Ala	Asp	Asn	Leu	Pro	Leu	Met	Val	Asp	Leu	
207			35					40					45				
209	ttg	gcc	gaa	tat	gaa	ggg	aaa	atc	gat	ctg	atc	tac	gcc	gat	ccc	cct	192
210	Leu	Ala	Glu	Tyr	Glu	Gly	Lys	Ile	Asp	Leu	Ile	Tyr	Ala	Asp	Pro	Pro	
211		50				55					60						
213	ttt	ttt	acg	gat	cgt	act	tat	gcg	gcg	cga	att	ggt	cat	ggg	gag	gat	240
214	Phe	Phe	Thr	Asp	Arg	Thr	Tyr	Ala	Ala	Arg	Ile	Gly	His	Gly	Glu	Asp	
215	65					70				75					80		
217	tcg	cgt	cgt	cca	caa	acc	tgg	cag	ctt	gca	gaa	gga	tat	acg	gac	gag	288
218	Ser	Arg	Arg	Pro	Gln	Thr	Trp	Gln	Leu	Ala	Glu	Gly	Tyr	Thr	Asp	Glu	
219				85					90					95			
221	tgg	aag	gat	tta	gat	gaa	tac	ctg	gac	ttc	ctt	tat	cca	cgc	ctg	gta	336
222	Trp	Lys	Asp	Leu	Asp	Glu	Tyr	Leu	Asp	Phe	Leu	Tyr	Pro	Arg	Leu	Val	
223				100					105					110			
225	ctg	atg	tat	cga	ctg	ctg	gca	cca	cac	gga	acg	ctc	tac	ttg	cac	ctg	384
226	Leu	Met	Tyr	Arg	Leu	Leu	Ala	Pro	His	Gly	Thr	Leu	Tyr	Leu	His	Leu	
227			115					120					125				
229	gac	tgg	cac	gcc	aat	gcc	tac	gta	cgt	gta	ctg	ctt	gat	gag	atc	ttc	432
230	Asp	Trp	His	Ala	Asn	Ala	Tyr	Val	Arg	Val	Leu	Leu	Asp	Glu	Ile	Phe	
231		130				135					140						
233	ggg	cga	cag	cgg	ttt	ctc	aac	gag	atc	gtc	tgg	atc	tat	cac	ggc	ccc	480
234	Gly	Arg	Gln	Arg	Phe	Leu	Asn	Glu	Ile	Val	Trp	Ile	Tyr	His	Gly	Pro	
235	145				150					155					160		
237	tca	gcc	atc	cga	cgc	gcc	ttc	aag	cgc	aaa	cat	gat	acc	atc	ttg	gtt	528
238	Ser	Ala	Ile	Arg	Arg	Ala	Phe	Lys	Arg	Lys	His	Asp	Thr	Ile	Leu	Val	
239				165					170					175			
241	tat	gtg	aaa	ggt	gaa	aac	tat	aca	ttc	aat	gcg	gat	gcg	gtt	cgt	caa	576
242	Tyr	Val	Lys	Gly	Glu	Asn	Tyr	Thr	Phe	Asn	Ala	Asp	Ala	Val	Arg	Gln	
243			180					185						190			
W--> 245	cct	tac	cat	ccg	agc	acn	cat	aag	acc	ttc	gct	tcc	tcc	ccg	aag	gcc	624
246	Pro	Tyr	His	Pro	Ser	Xaa	His	Lys	Thr	Phe	Ala	Ser	Ser	Pro	Lys	Ala	
247			195					200					205				
249	ggc	ttt	ggt	aag	gtg	ccg	gat	ctg	cag	cgc	ggc	aaa	gtg	ccc	gaa	gac	672
250	Gly	Phe	Gly	Lys	Val	Pro	Asp	Leu	Gln	Arg	Gly	Lys	Val	Pro	Glu	Asp	
251		210				215					220						
253	tgg	tgg	tat	ttt	ccg	gtc	gtg	gcc	cgt	cta	cac	cga	gaa	cgg	agc	ggc	720
254	Trp	Trp	Tyr	Phe	Pro	Val	Val	Ala	Arg	Leu	His	Arg	Glu	Arg	Ser	Gly	
255	225				230					235					240		
257	tat	ccg	act	caa	aag	cct	caa	gcc	ttg	ctg	gag	cgg	atc	ctg	ctg	gcc	768
258	Tyr	Pro	Thr	Gln	Lys	Pro	Gln	Ala	Leu	Leu	Glu	Arg	Ile	Leu	Leu	Ala	
259				245				250						255			
261	tcc	tcg	aac	gca	ggc	gat	ctg	gtg	gca	gac	ttc	ttc	tgc	ggc	tca	ggg	816
262	Ser	Ser	Asn	Ala	Gly	Asp	Leu	Val	Ala	Asp	Phe	Phe	Cys	Gly	Ser	Gly	

## RAW SEQUENCE LISTING

DATE: 08/05/2003

PATENT APPLICATION: US/09/689,343E

TIME: 18:28:00

Input Set : A:\PTO.AMC.txt

Output Set: N:\CRF4\08052003\I689343E.raw

```

263          260          265          270
265 aca acc gct gtg gtg gca gcc cgt ctg gga cgg cgc ttc ctg gtc aac      864
266 Thr Thr Ala Val Val Ala Ala Arg Leu Gly Arg Arg Phe Leu Val Asn
267          275          280          285
269 gat gca agc tgg cgc gcc gtt cat gtg aca cgc aca cgc ttg cta cgc      912
270 Asp Ala Ser Trp Arg Ala Val His Val Thr Arg Thr Arg Leu Leu Arg
271          290          295          300
273 gag gga gta agt ttc act ttt gaa cgc cag gaa act ttt act cta cct      960
274 Glu Gly Val Ser Phe Thr Phe Glu Arg Gln Glu Thr Phe Thr Leu Pro
275 305          310          315          320
277 atc cag cca ctt cca cca gat tgg ttg atc atc gcc gag gag cag att      1008
278 Ile Gln Pro Leu Pro Pro Asp Trp Leu Ile Ile Ala Glu Glu Gln Ile
279          325          330          335
281 cgc ctg caa gca ccc ttt ctg gta gat ttt tgg gaa gtg gac gat caa      1056
282 Arg Leu Gln Ala Pro Phe Leu Val Asp Phe Trp Glu Val Asp Asp Gln
283          340          345          350
285 tgg gat ggc aaa atc ttc cgc agc cgt cat caa ggc tta cgc tcc cgc      1104
286 Trp Asp Gly Lys Ile Phe Arg Ser Arg His Gln Gly Leu Arg Ser Arg
287          355          360          365
289 ctt cag gag cag gcg ccg ctg tct cta cca ttg acc ggg aat gga ctg      1152
290 Leu Gln Glu Gln Ala Pro Leu Ser Leu Pro Leu Thr Gly Asn Gly Leu
291          370          375          380
293 ttg tgt gta cgg gta gtg agc cgt gaa ggg gaa tac tat gag ttc aca      1200
294 Leu Cys Val Arg Val Ser Arg Glu Gly Glu Tyr Tyr Glu Phe Thr
295 385          390          395          400
297 ggt cga gcc gat agc cct cac ccc gta tcg ttt tga      1236
298 Gly Arg Ala Asp Ser Pro His Pro Val Ser Phe
299          405          410
301 <210> SEQ ID NO: 4
302 <211> LENGTH: 411
303 <212> TYPE: PRT
304 <213> ORGANISM: Unknown
306 <220> FEATURE:
307 <223> OTHER INFORMATION: Environmental DNA
309 <220> FEATURE:
310 <221> NAME/KEY: misc_feature
311 <222> LOCATION: (198)..(198)
312 <223> OTHER INFORMATION: Xaa = any amino acid
314 <400> SEQUENCE: 4
316 Met Pro Thr Leu Asp Trp Pro Gly Lys Gln Leu Ser Phe Pro Pro Ala
317 1          5          10          15
319 Thr Ser Leu His Leu Glu Ser Val Val Thr Glu Gly Ala Glu Ser Pro
320          20          25          30
322 Pro Asn Arg Leu Ile Trp Ala Asp Asn Leu Pro Leu Met Val Asp Leu
323          35          40          45
325 Leu Ala Glu Tyr Glu Gly Lys Ile Asp Leu Ile Tyr Ala Asp Pro Pro
326          50          55          60
328 Phe Phe Thr Asp Arg Thr Tyr Ala Ala Arg Ile Gly His Gly Glu Asp
329 65          70          75          80

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/09/689,343E

DATE: 08/05/2003  
TIME: 18:28:01

Input Set : A:\PTO.AMC.txt  
Output Set: N:\CRF4\08052003\I689343E.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:3; N Pos. 594  
Seq#:3; Xaa Pos. 198  
Seq#:4; Xaa Pos. 198  
Seq#:10; Xaa Pos. 19,24

**VERIFICATION SUMMARY**

DATE: 08/05/2003

PATENT APPLICATION: **US/09/689,343E**

TIME: 18:28:01

Input Set : **A:\PTO.AMC.txt**Output Set: **N:\CRF4\08052003\I689343E.raw**

L:31 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:1,Line#:29  
L:190 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:3,Line#:183  
L:245 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:576  
M:341 Repeated in SeqNo=3  
L:353 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:192  
L:409 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:5,Line#:407  
L:573 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:7,Line#:571  
L:709 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10 after pos.:16





1600

## RAW SEQUENCE LISTING

DATE: 08/05/2003

PATENT APPLICATION: US/09/689,343E

TIME: 08:19:32

Input Set : A:\Neb-181.txt

Output Set: N:\CRF4\08052003\I689343E.raw

3 <110> APPLICANT: Vaisvila, Romualdus  
 4 Morgan, Richard D.  
 5 Kucera, Rebecca B.  
 6 Claus, Toby B.  
 7 Raleigh, Elisabeth A.  
 9 <120> TITLE OF INVENTION: Method For Cloning And Producing The MseI Restriction  
 10 Endonuclease  
 12 <130> FILE REFERENCE: NEB-181  
 14 <140> CURRENT APPLICATION NUMBER: US 09/689,343E  
 15 <141> CURRENT FILING DATE: 2000-10-12  
 17 <160> NUMBER OF SEQ ID NOS: 21  
 19 <170> SOFTWARE: PatentIn version 3.1

## ERRORED SEQUENCES

172 <210> SEQ ID NO: 3  
 173 <211> LENGTH: 1236  
 174 <212> TYPE: DNA  
 175 <213> ORGANISM: Unknown  
 177 <220> FEATURE:  
 178 <223> OTHER INFORMATION: Environmental DNA  
 180 <220> FEATURE:  
 181 <221> NAME/KEY: CDS  
 182 <222> LOCATION: (1)..(1233)  
 183 <223> OTHER INFORMATION:  
 185 <220> FEATURE:  
 186 <221> NAME/KEY: misc\_feature  
 187 <222> LOCATION: (198)..(198)  
 188 <223> OTHER INFORMATION: Xaa = any amino acid

Does Not Comply  
 Corrected Diskette Needed

W--> 190 <220>  
 191 <221> NAME/KEY: misc\_feature  
 192 <222> LOCATION: (594)..(594)  
 193 <223> OTHER INFORMATION: N= G, A, C or T  
 195 <400> SEQUENCE: 3

W--&gt; 197

atg cct aca ctg gat tgg ccc ggt aaa cag tta agc ttc cca cca gct

48Met Pro Thr Leu Asp Trp

W--&gt; 198 1

5

10

15

E--&gt; 200 acc tcc ttg cat ctg gag agt gtg gtc act gag gga gcg gag tca ccg

96

201 Thr Ser Leu His Leu Glu Ser Val Val Thr Glu Gly Ala Glu Ser Pro

W--&gt; 202

20

25

30

E--&gt; 204 cct aat cgt ctg att tgg gcg gac aac ctg ccg cta atg gta gat ttg

144

205 Pro Asn Arg Leu Ile Trp Ala Asp Asn Leu Pro Leu Met Val Asp Leu

W--&gt; 206

35

40

45

insert  
 hard  
 return

## RAW SEQUENCE LISTING

DATE: 08/05/2003

PATENT APPLICATION: US/09/689,343E

TIME: 08:19:32

Input Set : A:\Neb-181.txt

Output Set: N:\CRF4\08052003\I689343E.raw

```

E--> 208 ttg gcc gaa tat gaa ggg aaa atc gat ctg atc tac gcc gat ccc cct      192
      209 Leu Ala Glu Tyr Glu Gly Lys Ile Asp Leu Ile Tyr Ala Asp Pro Pro
W--> 210      50                      55                      60
E--> 212 ttt ttt acg gat cgt act tat gcg gcg cga att ggt cat ggg gag gat      240
      213 Phe Phe Thr Asp Arg Thr Tyr Ala Ala Arg Ile Gly His Gly Glu Asp
W--> 214 65                      70                      75                      80
E--> 216 tcg cgt cgt cca caa acc tgg cag ctt gca gaa gga tat acg gac gag      288
      217 Ser Arg Arg Pro Gln Thr Trp Gln Leu Ala Glu Gly Tyr Thr Asp Glu
W--> 218      85                      90                      95
E--> 220 tgg aag gat tta gat gaa tac ctg gac ttc ctt tat cca cgc ctg gta      336
      221 Trp Lys Asp Leu Asp Glu Tyr Leu Asp Phe Leu Tyr Pro Arg Leu Val
W--> 222      100                      105                      110
E--> 224 ctg atg tat cga ctg ctg gca cca cac gga acg ctc tac ttg cac ctg      384
      225 Leu Met Tyr Arg Leu Leu Ala Pro His Gly Thr Leu Tyr Leu His Leu
W--> 226      115                      120                      125
E--> 228 gac tgg cac gcc aat gcc tac gta cgt gta ctg ctt gat gag atc ttc      432
      229 Asp Trp His Ala Asn Ala Tyr Val Arg Val Leu Leu Asp Glu Ile Phe
W--> 230      130                      135                      140
E--> 232 ggg cga cag cgg ttt ctc aac gag atc gtc tgg atc tat cac ggc ccc      480
      233 Gly Arg Gln Arg Phe Leu Asn Glu Ile Val Trp Ile Tyr His Gly Pro
W--> 234 145                      150                      155                      160
E--> 236 tca gcc atc cga cgc gcc ttc aag cgc aaa cat gat acc atc ttg gtt      528
      237 Ser Ala Ile Arg Arg Ala Phe Lys Arg Lys His Asp Thr Ile Leu Val
W--> 238      165                      170                      175
E--> 240 tat gtg aaa ggt gaa aac tat aca ttc aat gcg gat gcg gtt cgt caa      576
      241 Tyr Val Lys Gly Glu Asn Tyr Thr Phe Asn Ala Asp Ala Val Arg Gln
W--> 242      180                      185                      190
E--> 244 cct tac cat ccg agc acn cat aag acc ttc gct tcc tcc ccg aag gcc      624
      245 Pro Tyr His Pro Ser Xaa His Lys Thr Phe Ala Ser Ser Pro Lys Ala
W--> 246      195                      200                      205
E--> 248 ggc ttt ggt aag gtg ccg gat ctg cag cgc ggc aaa gtg ccc gaa gac      672
      249 Gly Phe Gly Lys Val Pro Asp Leu Gln Arg Gly Lys Val Pro Glu Asp
W--> 250      210                      215                      220
E--> 252 tgg tgg tat ttt ccg gtc gtg gcc cgt cta cac cga gaa cgg agc ggc      720
      253 Trp Trp Tyr Phe Pro Val Val Ala Arg Leu His Arg Glu Arg Ser Gly
W--> 254 225                      230                      235                      240
E--> 256 tat ccg act caa aag cct caa gcc ttg ctg gag cgg atc ctg ctg gcc      768
      257 Tyr Pro Thr Gln Lys Pro Gln Ala Leu Leu Glu Arg Ile Leu Leu Ala
W--> 258      245                      250                      255
E--> 260 tcc tcg aac gca ggc gat ctg gtg gca gac ttc ttc tgc ggc tca ggg      816
      261 Ser Ser Asn Ala Gly Asp Leu Val Ala Asp Phe Phe Cys Gly Ser Gly
W--> 262      260                      265                      270
E--> 264 aca acc gct gtg gtg gca gcc cgt ctg gga cgg cgc ttc ctg gtc aac      864
      265 Thr Thr Ala Val Val Ala Ala Arg Leu Gly Arg Arg Phe Leu Val Asn
W--> 266      275                      280                      285
E--> 268 gat gca agc tgg cgc gcc gtt cat gtg aca cgc aca cgc ttg cta cgc      912
      269 Asp Ala Ser Trp Arg Ala Val His Val Thr Arg Thr Arg Leu Leu Arg
W--> 270      290                      295                      300
E--> 272 gag gga gta agt ttc act ttt gaa cgc cag gaa act ttt act cta cct      960

```

## RAW SEQUENCE LISTING

DATE: 08/05/2003

PATENT APPLICATION: US/09/689,343E

TIME: 08:19:32

Input Set : A:\Neb-181.txt

Output Set: N:\CRF4\08052003\I689343E.raw

```

      273 Glu Gly Val Ser Phe Thr Phe Glu Arg Gln Glu Thr Phe Thr Leu Pro
W--> 274 305                310                315                320
E--> 276 atc cag cca ctt cca cca gat tgg ttg atc atc gcc gag gag cag att      1008
      277 Ile Gln Pro Leu Pro Pro Asp Trp Leu Ile Ile Ala Glu Glu Gln Ile
W--> 278                325                330                335
E--> 280 cgc ctc caa gca ccc ttt ctc gta gat ttt tgg gaa gtg gac gat caa      1056
      281 Arg Leu Gln Ala Pro Phe Leu Val Asp Phe Trp Glu Val Asp Asp Gln
W--> 282                340                345                350
E--> 284 tgg gat ggc aaa atc ttc cgc agc cgt cat caa ggc tta cgc tcc cgc      1104
      285 Trp Asp Gly Lys Ile Phe Arg Ser Arg His Gln Gly Leu Arg Ser Arg
W--> 286                355                360                365
E--> 288 ctt cag gag cag gcg ccg ctc tct cta cca ttg acc ggg aat gga ctg      1152
      289 Leu Gln Glu Gln Ala Pro Leu Ser Leu Pro Leu Thr Gly Asn Gly Leu
W--> 290                370                375                380
E--> 292 ttg tgt gta cgg gta gtg agc cgt gaa ggg gaa tac tat gag ttc aca      1200
      293 Leu Cys Val Arg Val Val Ser Arg Glu Gly Glu Tyr Tyr Glu Phe Thr
W--> 294 385                390                395                400
E--> 296 ggt cga gcc gat agc cct cac ccc gta tcg ttt tga      1236
      297 Gly Arg Ala Asp Ser Pro His Pro Val Ser Phe
E--> 298                405                410

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/09/689,343E

DATE: 08/05/2003  
TIME: 08:19:33

Input Set : A:\Neb-181.txt  
Output Set: N:\CRF4\08052003\I689343E.raw

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:3; Line(s) 197

Use of <220> Feature (NEW RULES) :

Sequence(s) are missing the <220> Feature and associated headings.

Use of <220> to <223> is MANDATORY if <213> ORGANISM is "Artificial Sequence" or "Unknown". Please explain source of genetic material in <220> to <223> section (See "Federal Register," 6/01/98, Vol. 63, No. 104, pp.29631-32) (Sec.1.823 of new Rules)

Seq#:11

**VERIFICATION SUMMARY**

DATE: 08/05/2003

PATENT APPLICATION: **US/09/689,343E**

TIME: 08:19:33

Input Set : **A:\Neb-181.txt**Output Set: **N:\CRF4\08052003\I689343E.raw**

L:31 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:1,Line#:29  
L:190 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:3,Line#:183  
L:197 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:17  
L:198 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
L:200 M:254 E: No. of Bases conflict, LENGTH:Input:96 Counted:48 SEQ:3  
L:202 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
M:254 Repeated in SeqNo=3  
L:206 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
L:210 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
L:214 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
L:218 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
L:222 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
L:226 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
L:230 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
L:234 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
L:238 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
L:242 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
L:244 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:528  
M:341 Repeated in SeqNo=3  
L:246 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
L:250 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
L:254 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
L:258 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
L:262 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
L:266 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
L:270 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
L:274 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
L:278 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
L:282 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
L:286 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
L:290 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
L:294 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
L:298 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3  
L:298 M:252 E: No. of Seq. differs, <211> LENGTH:Input:1236 Found:1188 SEQ:3  
L:352 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:192  
L:408 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:5,Line#:406  
L:572 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:7,Line#:570  
L:708 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10 after pos.:16  
L:717 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ#:11, <213>  
ORGANISM:unknown<220><223> Primer MseI-IP1  
L:717 M:258 W: Mandatory Feature missing, <223> Tag not found for SEQ#:11, <213>  
ORGANISM:unknown<220><223> Primer MseI-IP1  
L:717 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:11,Line#:717